

知識社会時代の到来と 新しい知識創造検索システムの構築

三輪 晴 治

The New Search System for the Information Age at the
Beginning of the Twenty-first Century

Seiji MIWA

ABSTRACT

The current keyword type search system on websites has a lot of problems. Current consumer oriented search strategies use keywords to access web pages ranked by page popularity. The searcher are increasingly frustrated by an inability to efficiently and precisely discover and retrieve the information that they seek, because the current search system does not understand the meaning of the keywords.

We need to develop semantic maps of source documents and rank and group retrieved documents according to how well they match semantic maps of user queries and how well they fit user-specified knowledge contents.

The real information age and the Soft bath age will come after such relevant search systems are developed.

KEYWORDS : Search Economy, Taxonomy, Structured Knowledge, Ontological Search Explorative Search, Knowledge of Crowd, Reckless Drivers of Crowd.

はじめに

これまでの経済社会は、「資源・素材」を獲得し、それをいろいろと加工して商品を生産し、販売するという世界であった。しかし20世紀の後半から、資源・素材に加えて「知識・情報」を獲得し、それを加工して素材の中に注入したり、あるいは知識・情報そのものを加工し、組み合わせて知識商品・ソフトを生産、販売する時代になってきた。これが「知識社会」の基本的な構造で、知識・情報で富を創造し、獲得する社会である。

これまでの経済社会では、知識・情報は偏在して、ある一部の人、企業が所有し、情報の非対象性という独占力で優位なビジネスを展開していた

が、その知識・情報がウェブ上にオープンになり、情報の非対象性が崩れてきつつある。これが知識社会の展開を進める大きなエネルギーになっている。

その知識・情報が、1990年代からインターネットで作りあげられたウェブのなかに存在するようになった。以来このウェブにおける知識、情報が猛烈な勢いで増殖し続けている。アメリカの調査会社の IDC によると、2006年に世界で作られたデジタル情報は1610億ギガバイトにのぼるという。これは過去に印刷されたすべての本の300万倍の情報量に相当し、2010年に作られる情報は、その6倍以上になると予測している。アメリカのカリフォルニア大バークレー校の調査によると、世界で毎年生み出されている情報の量は、人類が2000年かけて蓄積してきた情報の量を超えているという。

他の推定では、現在世界のウェブには800億の

2007年9月20日受付, 2007年12月13日最終受付
三輪晴治 四国大学経営情報学部経営情報学科客員教授・イーエイシック・ジャパン株式会社
Seiji MIWA, Member(eASIC JAPAN Corporation, 3-17-6 Shinyokohama, Kohoku-ku, Yokohama-City, Kanagawa, 222-0033 Japan).
四国大学経営情報研究所年報 No. 13 pp.35-53 2007年12月

ホームページがあるという。『グーグル・アマゾン化する社会』森健 光文社新書)

いずれにしても、この数はこれからどのくらいのスピードで増殖するのか計り知れない。

人間の歴史での情報の増大は、文字、図形が創造されて、それがパピルスに記録され、洞窟に描かれた壁画、さらに紙の発明とともにグーテンベルグの印刷機で書物がどんどん出版されるようになり、そして近年のタイプライターの出現により個人レベルからも、膨大な情報が増殖、蓄積されてきた。同時に人間の誕生以来、肉体のなかに記録されてきた「DNA という情報」も解読されるようになり、人間世界の情報の増大に拍車をかけている。

その情報の急速な増殖は、半導体の記憶素子の発明で、新しいフェーズにはいった。つまり1980年代の半ばからの世界の半導体の記録素子(DRAM)の生産数の伸びを見ると、その情報の急速な驚異的な増大が分かる。1985年からDRAMの総ビット数は、対数表軸の中で、直線で伸びており、それは今日でも続いていることに気付く。10の18乗、20乗と増殖している。これは大変なことである。それがインターネット、ウェブの世界に乗り始めて、さらにその増殖のスピードに拍車がかかっている。DRAMとコンピュータが知的生産の新しいフェーズをもたらしたことになる。

そこでその膨大な「知識・情報」を、個人、企業がどのように効果的に入手、確保することが出来るかが重要になってきた。特に膨大な量の、しかも刻々増殖、拡大しているウェブ上の知識・情報をいかに効率よく的確に見つけ、獲得するかが重要になり、さもないと情報の洪水のなかで情報の欠乏という現象が起こる。つまり求める情報、知識をいかに効果的に「検索する」かである。サーチ・エコノミー時代といわれる所以である。

現在のウェブ上の情報・知識の検索の仕組みは欧米の企業のものが主流になって、使われている。これからはじまるウェブでの検索クエリーによる「サーチエコノミー社会」が本格的に進行するな

かで、今日欧米諸国のリードする検索システムが急速に普及し始めているが、残念ながら今日の情報検索システムにはフラストレーションが鬱積しているといわれている。

欧米の諸国は、この情報検索システムは情報技術の開発力、国の安全保障力にも大きな影響があると認識し、そのために国家プロジェクトとして新しい検索システムの開発に動き始めている。

フランスは昨年(2006)1月に新検索エンジン「クエロ(Quaero)」の開発を国家プロジェクトとして進めると宣言し、約400億円の投資を決定している。

現在の世界の検索システムはアメリカのグーグルが支配しているが、そのアメリカでも、国防省が更に新しいコンセプトの検索システムを開発しはじめているという。

これからの構築されるべき新しい情報検索システムはどんなものでなければならないか。これが最近急に重要な課題として世界的に注目を浴びるものとなった。

「サーチエコノミー社会」は、「ソフト・パス」というコンセプトでも表現されているが、本格的に展開する21世紀の『知識社会』において、世界の誰もが平等に、地球上に散在し、また猛烈な勢いで増殖を続ける知識、情報にアクセスし、検索して、それを加工し、組み合わせて新しい知識商品を創り、イノベーションを進める仕組みが確立されなければならない。

そのためには効果的に、的確に求めるものを検索して結果を得なければならない。しかも単に探したいものをキーワードで検索して、その検索結果を受け取るということだけではなくて、その人の連想力、ひらめきをもとに新しい知的創造活動をすることを助けるものでなければならない。ここに新しい検索システムの課題がある。

これはまた21世紀の国力という意味での「ソフト・パワー」としての国の「情報力」に繋がるものである。これを可能にするような「情報知識検索システム」はどのようなものにならないか21世紀の知識社会の構造を構築するう

えでの大きな課題であり、この基本的考え方をここで模索してみたい。この点に関して、グローバル世界での、地域国家社会としての日本はどのようなことをしなければならないかを考えてみる。

(1) 知識社会の意味

「知識社会」、「情報化社会」の到来と叫ばれはじめてからすでに久しい。工業化社会と比べて、単に情報技術を利用するということだけではなく、富の創造、獲得が、「物理的な資源を獲得し、加工」して商品を創るというプロセスではなく、複写可能な「情報、知識」を創造し、加工し、組み合わせて「商品」を創るプロセスに変わってきた。つまりこれまでの「資源の獲得競争」から、「情報の獲得」「知識の創造競争」に変わってきたのである。

しかしこの「情報商品」は近年になって初めてでてきたものではなく、資本主義経済の発生以来存在するのであり、ハードの商品のなかにも「情報、知識」が注入、転写され、どんどん侵入してきている。つまり付加価値として「情報・知識」の部分が時代とともにだんだん大きくなってきたのであり、それが1990年以降の半導体産業の進展、コンピュータ・コミュニケーションの時代に入り、質的転換を遂げてきたのである。

別の言い方をすれば、基本的には、資本主義経済の商品が、人間が物理的に生きていくための「生活必需品」から人間の感情、好み、趣向を満足するような「便益品」の時代になり、その商品のなかの情報、知識の占める割合がますます大きくなってきていることである。そして「生活必需品」の需要の充足をはたし、「便益品」としての「情報商品」の量、比率が20世紀後半にはいり急速に拡大してきたということである。

とくに「情報、知識」そのものが、インターネット、ウェブ上に乗り、動きまわり、増殖しはじめた1990年から、これまでとは違った新しい情報化社会、知識社会が到来することになった。つまり膨大な情報、知識がウェブ上でオープンになって

きたことにより、情報の独占で成り立ってきたこれまでの社会が大きく変貌しているのである。

これまでの工業社会では生産手段は資本家が所有し、それを雇われた労働者が使うという形をとっていたものが、この新しい時代では、知的生産手段は知的労働者が所有し、労働者自身がオープンなウェブ上の知識、情報、データを検索し、加工し、組み合わせてイノベーションを起こす時代になろうとしている。

(2) 「ソフト・バス」の世界

来るべき情報化社会、知識社会の経済社会構造をさらに考えてみよう。

「第1量子化」の理論では、素粒子の増大は、情報の増大として、エントロピーの増大になり、無秩序化が進むことになる。社会も情報が無秩序に増殖すると社会の秩序は崩壊し、アナーキーな社会になる。現在はその無秩序な現象が起こっているといえよう。

しかし物理の世界で、その素粒子の増大が進み、「第2量子化」理論に移る。素粒子が増大して、「場」の概念のもとに、無秩序に存在するのではなく、素粒子はお互いに非常に強い力で相互作用を持ちながら動きまわっているのだということになる。これが「第2量子化」の中心理論である。しかしその相互作用があまりに強くなるとその系全体は不安定になる。やがては相移転の現象となる。これを、物理学では、ある一つの系の周りには、必ず大きな熱量を持つ「ヒート・バス(熱浴)」がついていると考える。つまり熱平衡状態にある大きな系(熱浴)と接触している。

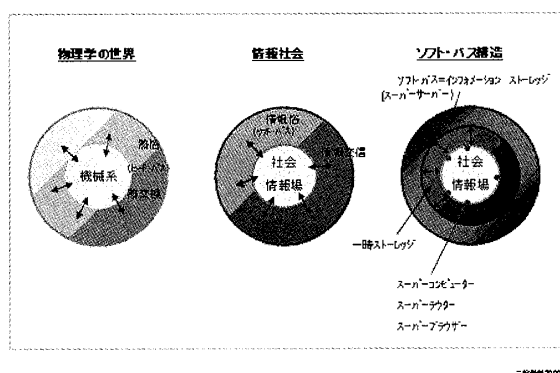
このヒート・バスが熱の授受をしてくれるので、ある機械系のメカニズムはその機能を混乱も破壊も無く、展開できるのである。電気モーター、ガソリンエンジンを考えても、放熱板、ラジエーターを装備しなければ、駆動を続けることはできない。

物理の世界のアナロジーを情報の世界に適用するとこういうことになる。

世界のウェブ上で増殖する知識を検索し、獲得

し、加工し、組み立て「知識商品」を造り、販売するには、人類の蓄積してきた、またこれから創り上げる情報、知識を、人間社会に適切に授受する「ソフト・バス（情報浴）」がなければならない。これがなければ情報化社会は無秩序となり、争いの場となり、崩壊する。これが田中 昭二氏の提唱されている「ソフト・バス」の概念である。

ソフト・バスの世界



（図-1）

経済学者のフリードリッヒ・ハイエクは、資本主義の市場、メカニズムを「知識」という側面から解明しようとした。彼は「知識」の分業という概念で、資本主義社会の市場活動を情報システムとして捉えた。市場は競争のもとに、分散した個々の知識を交換、結合して商品生産、販売、消費をする場であるとした。これを「カタラクシー」と呼んだ。

ハイエクは、情報、知識は万人に平等に保持されていないために、その必要な情報、知識を獲得する競争が生まれ、経済社会が維持され、間断なきイノベーションが起こるとした。つまりハイエクは情報の非対称性を前提にしていた。

これからの情報化社会では、一般大衆がオープンになった膨大な情報、知識のなかから、自分に意味のある情報、知識を検索し、それを「新結合」してイノベーションを進め、新しい商品を創り、販売、消費する世界である。これはあたかもヒート・バスの世界で「機械系のシステムに熱を授受する」ように、「経済活動で知識労働者に知識・情報を授受する」仕組みである。

つまりこれは、情報処理技術と通信技術のアップヘーベンが、新しい情報場としての「ソフト・バス」を作り上げることになる。

（参照：『第2情報化社会の考察』田中 昭二
Nikkei Electronics 1996. 7.15）

（3）情報化社会の新しい構造の兆候

この新しい社会の特徴的な構造は、希少な資源に依存する経済から、基本的には無尽蔵の、しかもコピーがいくらでもできる資源である「情報、知識」を経済発展の主因とする知識商品経済へとかわった。その新しい社会の骨格の兆候がすでに現れてきている。

3-1）知的生産手段を所有する知的労働者

新しい知識社会では、知識労働者は自分の生産手段をもつことになる。しかしその生産手段のインフラができなければならない。これが「ソフト・バス」である。

これまで資本家が生産手段を所有し、労働者はその資本家の所有する生産手段で物理的資源を加工することに賃労働をしている構造から、知識商品の生産手段を自分で持ち、自分の力でウェブ上にある情報、知識の素材を検索、収集して新しい自分の商品を生産する時代に突入してきているのである。生産手段としてのパソコン、携帯端末これからこれに家電機器も入ってくる。知識労働者が自分の持つパソコン、携帯情報端末で、社会のインフラである「ソフト・バス」上で、知識商品を創り、イノベーションを進める社会となる。

特に最近のウェブの世界で、「こちら側と向こう側」という概念ができ、こうしたソフト・バスの環境が現実のものとなってきている。

「従来のウェブとの最大の違いは、パソコンやケータイなどの端末上ではなくウェブの中で情報を手軽に操作できることだろう。わかりやすい例はウェブメール。これまでのように端末同士でメールを交換するのではなく、サーバー上で直接メールを読み書きしてしまう」（日本経済新聞

西垣 通 『今を読み解く』 2006年7月2日)

つまり自分の持っているパソコン、携帯情報端末は軽いものになるが、膨大な数のPCを繋げたスーパーコンピュータ、スーパーラウター、スーパーブラウザー、スーパータクソノミー、膨大なデータベースからなる「ソフト・バス」をあたかも自分のものとして駆使することができる。こうした環境がますます一般大衆の知識労働活動の大きな流れを加速している。

最近ではウェブ上の一般検索、Wikipediaその他で、これまでの重いハードコピーの百科事典、業界誌、学術誌、統計資料からではなく、ウェブから知りたい情報、知識を簡単に引き出し始めている。しかし現実には今日の欧米企業のドライブするウェブでの情報・知識検索結果には、多くの人がフラストレーションを感じていることも事実である。

しかしここで重要なことは、時々刻々増殖を続ける知識・情報の量が増大し、知識労働者の生産性に情報のオーバーロードが障害を与え始めているという指摘である。これからの頭脳労働者が常時「的確な情報、知識」をウェブから取り寄せて、それを加工、組み合わせて知的イノベーションを進める環境を作る必要がある。本当の意味でのウェブでの商取引は、もう少し検索結果が信頼できる仕組みにならなければ、起こらないという意見がある。

3-2) 企業の有形固定資産と無形固定資産の比率

既に知識、情報、ソフトが、人間社会が創り上げる商品の付加価値のますます大きなものになっており、経済社会の資産の中心となる。

「ブルッキングズ研究所の調査によれば、1982年にはすでに、鉱工業企業ですら時価ベースの総資産のうち、無形資産が38パーセントを占めていた。その10年後というと、ドット・コム boom と崩壊のはるか前だが、それでも無形資産が62パーセントと総資産価値の3分の2近くを占めていた。」(『富の未来 下』アルビン・トフラー、ハイジ・トフラー 講談社 P.100)

「今日、スタンダード・アンド・プアーズ (S & P) の株価指数「S&P500」に名を連ねる企業の市場価値の約85%は無形資産によるものだ。これを敷延すれば、アメリカの輸出市場の大部分が無形資産に負っていることになる。」(『Innovate America』)

今日の自動車でも電子機器、ソフトによる付加価値は、全体の60%以上になっており、その比率は今後さらに増大するといわれている。自動車の価値のなかで、物理的なエンジン、トランスミッション、ボディの価値はだんだん相対的に低下してくる。

しかしまだソフト、知識、情報の価値を評価する適切なメソッドロジーが十分確立していないので、今日の無形資産は過小評価されていると言わなければならない。

また企業の持てる知的財産をフルに活用すれば今の利益の3倍は可能になるといわれているように、また知的財産をフルに活用するインフラができていない。

3-3) 科学・技術の開発と情報検索

これまでの工業社会での巨大企業がドライブしてきた「リニアー方式」の研究開発が衰退し、新しい開発の仕組みが出来つつある。つまりかつてのナイロンの発明、ラッカーの発明など、個人の思い違いによる偶然の発明というリニアーモデルで研究開発、商品化ではなくなっていった。新しい開発モデルは、1980年からシリコンバレーがパーソナル・コンピュータ、ネットワークングで作り上げた産業・アントルプルヌール・クラスターによる、多くの知識の結合、融合のプロセスである。こうした新しいイノベーションのプロセスが出来上がりつつある。これがシュンペーターのいうイノベーションということになる。

いうまでもなくこれには世界のソフト、知識、情報の結合、新結合の新しい社会的な仕組みが必要になるが、幸いにもウェブの世界がこれをバックアップしてくれている。今日のウェブ上の情報、知識を検索すれば、ミサイルの製造法もサリンの

製造法も分かるといわれているように、膨大なデータ、情報、知識が存在する。しかもこの情報、知識は時々刻々どんどん増殖している。

つまりこれからの科学技術開発も、世界のウェブ上のもの、それからリンクされたデータ、情報、知識から検索して適切なものを結合するプロセスになってきている。これがイノベーションに繋がる。コンビナトリー・デベロップメントとも繋がる。さらに消費者も商品開発に参画しており、ウェブでこれが現実のものとなりつつある。アメリカの21世紀の戦略である「イノバート・アメリカ」はこれを明確に目指している。

3-4) 多品種少量の生産消費の世界

ワン・ツー・ワン・ビジネス、プロシューマーの時代と言われている。知的財産、ソフトはコピーがいくらでも出来るという性格から、少量生産でもコストは高くない。「ロングテール」の中から、グローバル市場でのビジネスが展開できるようになる。

インターネット、ウェブの世界ではロングテールとして表現されている市場の分野にも効果的な商品開発、供給が出来る。ナローキャストが出来る。

この世界で始めて筆者のいう真の『カスタメーション』（カスタム+オペレーション 『創造的破壊』 三輪 晴治 中公新書）が誕生し、経済社会構造が大きく変貌することになるだろう。

こうしたことがこれからインターネットとウェブサイトによりますます進められ、まだ我々の頭では考え付かない新しい世界とビジネスモデルが展開されることになる。

あと100年たてば資本主義経済のコンセプトも大きく変貌することになるかもしれない。

しかしこのような良いことづくめではない。情報というこれからの知識社会での人間生活の最も基本となるものがある国に所有され、支配されるということは憂慮すべきことである。

つまりこうした知識社会での「ソフト・バス」というコンセプトの国民大衆にオープン化される

知識データベースの性格を深く考えて見なければならぬ。

モノとは違って情報はすぐ人に伝わり、コピーされるものであるが、そのデータベースに、あるいはその検索システムに人為的な恣意が入ってはならないことである。かつての「ドイツ参謀本部の情報戦略」では、敵の行動計画は、一般に誰でも入手できる、公開されている新聞、雑誌、刊行物をつぶさに解読するとそれが理解できるとしていた。つまり誰でもアクセスできる公開の情報を構造的に分析すると、世の中の殆どの動き、内容が理解できるということである。そうした意味ではこのウェブ上の情報はその公平性と、構造化が極めて重要である。

近代科学の進歩は、分野の細分化による専門化するプロセスであった。狭い分野を深く掘り下げて科学、技術を探求していくことが高度に進んできた。しかし最近の技術、商品、産業、社会現象はますます複雑なものになり、あらゆる分野でグローバル化現象が進行してきており、人間が関知し、認識する局面もますます複雑化してきている。最近のプロジェクト、商品システムは膨大な数の要素が組み合わされ、それが互いに関連し、コミュニケーションして出来ている。特に環境問題は地球規模のスケープで問題が繋がりを、それを解明し、ソリューションを確立することが急務となっている。これまでのように人間の直感ではもはや解決できないものとなった。そうした意味ではこれまで細分化を続けてきた膨大な科学技術知識を意味的に関連付け、それを構造化しなければならない段階になったということである。つまりものごとの全体像を描いて共有することであるが、今そうした全体像を描く力が不足していることが問題として浮かび上がっている。そこで叫ばれてきているのが「知の構造化」である。

20世紀は「科学の黄金時代」であったが、21世紀は「科学統合の時代」といえよう。

かつて1750年ころから20年をかけてドゥニ・ディドロとジャン・ダランベールが進めたのも「知識の構造化」であった。百科全書派と呼ばれ

るものであるが、知識と知識を意味的に繋げて体系的に整理しておくことで新しい知識が生まれるとされている。

つまり知識を意味的に関連付けて構造化し、クロスリファレンスを含めて、科学と価値を繋げることである。しかし当時の知識の量と複雑度はさほどでもなく、人間の頭の中での処理で何とかでき、これ以上進展しなかった。

しかし今日の知識の量と複雑度は、科学の発達の段階の違いでこれまでとは比較にならない膨大な量と複雑度をもってきており、しかもそれが物凄い勢いで増殖していることである。

その基礎の一つとなるのがこうしたウェブ上での知識の検索システムである。

この知識の検索システムを、求める本当の知識の獲得が可能になるようなものを確立しなければならない。それにはデータベースとして知識の構造化をしなければならないことになる。

そして情報操作の問題を含めて、グローバル社会では、各国があるいは地域社会がコンピュータ、サーバー、ルーターを持ち、データの分散、独立管理ができる状態にする必要がある。

いずれにしても1995年ぐらいから、こうした知識社会の時代に入りアメリカその他の国は、その産業経済の構造を大きく変えてきている。しかし日本は依然として旧来の産業構造を踏襲していることが問題になっている。

(4) ウェブでの情報・知識検索システムの現状と問題点

4-1) 今日の検索システムの問題点

これからの知識社会では、国民の誰もが公平に、平等にウェブの世界に散在する情報、データ、知識にアクセスして、それを加工し、組み合わせて自分の求める新しい知識商品（ハードのもののなかに注入・転写するものを含む）を創り出す仕組みを確立する必要がある。情報、知識は、それがパブリックになったとたん市場価値が減衰し、なくなるというものであるから、知識社会で新しい

価値を創造し、提供できるためには、常にアップデートな、世界に散在する知識、情報、データを加工、組み合わせて新しい知識として創造することを可能にする環境を造る必要がある。これが21世紀の知識社会における創造活動である。

これからの知識社会は、自分の求める知識、情報を効果的に検索することができると同時に、イノベーションを進めるために、知識の創造性を追求する上でどのような仕組みにならなければならないかを考えなければならない。

しかし近年のウェブ上での検索システムが急速拡大してきた中で、情報、知識の入手のコストが格段に下がってきたことがこのソフト・バスの社会に進む大きなイナーシャーになっているといえる。特にグーグルその他の検索システムは情報、知識の検索を基本的には無料で提供しているからである。グーグルは検索連動型広告で収入を得て、情報の検索を無料にするビジネスモデルにより大きな成功を成し遂げている。

これからの新しい情報検索システムについての第一の要件として、自分が欲しいものを、意味的に関連するものを含めて検索できることが必要である。しかし逆に自分の求めるものでないものが、つまり関係ないものが検索結果として出てくるのでは困る。またさらに、実際にはその検索サイトのデータベースの中に自分の欲しい情報、知識が存在するにも拘らず、検索結果に出てこないことも、言うまでも無く、望ましくない。

さらに第二の重要な要件は、前述のとおり、知識社会での創造的活動では、前もって欲しいと思った「もの」を検索結果で得ることでは十分でなく、人間の思考過程で起こる「考えの発展」、「創発」といわれる飛躍」が起こる環境が必要である。今日の検索システムはフラストレーションが起こり、楽しくないという声が上がっているものこの問題である。予め想定した「情報・知識」を文字・図形の形で確認するのではなく、予め想定したものを越えた新しい発見、飛躍としての情報・知識が出てくる必要がある。これが「知識社会での創造プロセス」である。

ある人は、インターネットの世界の「偶有性」という言葉で表現しているように、半ば偶然的に、半ば必然的に新しい創造が生まれてくる性質があり、これを検索システムのなかでさらに追求しなければならない。

4-2) キーワード型検索システム

現代の欧米企業のドライブするウェブ上での検索システムで、一般的にとられているのは「キーワード型検索」システムである。

「キーワード型検索システム」とは、検索サイトの画面上のサーチボックスに「キーワード」を入力して、求めるものを検索するシステムである。検索対象となるキーワードは、クローラーあるいはスパイダーと呼ばれるソフトでウェブに張り巡らされたリンクで世界のウェブページを探しまわり自動的にキーワードを抽出するものと、各ウェブページのオーナーが検索サイトに依頼してキーワードを登録するものとがある。

この二つのキーワード型検索システムに共通の問題点は、検索サイトに登録されたキーワードと情報を探している者によって入力されるキーワードの文字的に一致するものしか検索結果に出さないということである。つまり情報・知識を探している者はある検索サイトに登録されているキーワードを予め知っているわけではないので、その検索サイトに登録されたキーワードが一致しなければ結果として出てこない。キーワード型検索システムは、自分がどんな情報を欲しいかを明確に理解していない者には、しかもそのキーワードが検索サイトに登録されていると確信のできない者には、適切な検索が出来ないという基本的に矛盾したシステムであることが分かる。

もう一つの問題点は、前述のように、キーワード型検索は、入力されたキーワードという「文字」に、検索サイトの登録されたキーワードという「文字」と一致したものを検索結果として提示するものであり、キーワードの「意味的」な関連において検索サイトにある情報を結びつける機能を持っていないということである。ある情報を求める者

は、入力するキーワードと意味的に関連したものを探しているのである。

更にもう一つの問題点は、世の中が猛烈な勢いで変化していることである。これは世界の製品、サービス、企業名、商品名、事象が刻々変化し、それに関連して新しい言葉が生まれ、その意味的関連が変化することにどう対応するかである。

もう少し具体的に言えば、サーチエンジンに入力されたキーワードは『言語』ではない。それは単なる『カタカナ、平仮名、アルファベット、あるいは漢字』の文字の繋がったものであるということを理解しなければならない。つまり今日の検索システムにおけるコンピュータは、その羅列された文字の「意味」を認識し、特定はしない。コンピュータは単に入力された文字の羅列とキーワードとして登録されたものとマッチしたものを確認し、そのキーワードの存在する情報・書類、サイトを引き出すだけである。

ある人は言う。藁をもつかむ思いで情報を探している人は、無理やりにキーワードをひねり出し検索画面に入力するが、その結果、入力したキーワードの全部または一部を含んではいるが、求めるものとは関係のない膨大な情報のリストに出くわす。あるいは、その入力キーワードに該当する情報はゼロでしたという結果になる。野山に野生のマツタケを探しに出かけるようなもので、運がよければ本物のマツタケに出会えるが、運が悪ければ毒キノコになるし、あるいは全く収穫なしということになる。

つまり、現実には、こうしたキーワード型検索システムでは、あるキーワードを入力すると、そのキーワードでタグをつけられた膨大な数の情報・書類が引き出される。この膨大な情報・書類が引き出されても、どれが求めるものかを判断するのは極めて困難であるので、グーグルはPageRankというアルゴリズムを開発した。このビジネスモデルが、新しいウェブ広告産業を創り上げたのである。

グーグル社の特許になるというものだが、検索結果の表示リストの順位を、いかに多くの人がそ

の情報をクリックしたか、あるいはいかにより多くの人がそれを引用し、リンクさせているかを基準にして、検索された膨大なものの中から順次並べるというものである。

ケインズ経済理論で「美人投票」という原理がある。実際にイギリスであったことであるが、投票者が女性の100枚の写真の中から最も容姿の美しい6枚を選び、その選択が投票者全体の平均的な好みに最も近かった人に賞品があたえられる。これが資本主義経済の動きは、「望ましくないことである」が、この美人投票に似ているとケインズはいった。「誰が本当に美人であるか」ではなく、「如何に多くの人が誰が美人と思っているか」が中心の点となっているのである。「皆で渡れば怖くない」という考えである。これが資本主義経済における「市場の暴走、市場の失敗」といわれるものに繋がる。

この資本主義経済の「好ましくない性質」にグーグルの検索は似ているといえる。つまり何を検索ユーザーが求めているかという点では、必ずしも正しい検索結果が出てくるとは限らないのがこのシステムである。経済の場合は、多くの人が投票したものが、社会、企業、大衆にとって好ましくないものであると後でわかると、それに対して反動が起こり、詐欺的なものであれば社会的に処罰されることになり、事実現の経済活動の世界ではそれが起こってきているのである。

「群衆の叡智」とは聞こえは良いが、「群衆の暴走」になる可能性もあり、情報の操作の危険性をもっていることも事実である。

このようにグーグルは、この PageRank というコンセプトのアルゴリズムで、ビジネスの収入を図っている。つまり PageRank のコンセプトで、情報のオーナーがそのクリック、リンクの量に応じて広告料をグーグルに支払うという「検索連動型広告」がビジネスモデルである。しかし、この広告の支払い主が、「自分は本当は美人ではないのでは」と疑問を抱き始めると、このビジネスモデルはやがて崩れることになる。

ある見方では、現在のグーグルのような検索連

動型広告のキーワード型検索システムでは、その検索結果は、あらかじめグーグルと契約をしたサイトのものを出してくるという仕組みになっている。つまり契約して広告を払うサイトのものでなければ出てこないことが多いといわれている。ユーザー全体が広告ビジネスの仕事を無意識のうちにさせられているのであり、それをまたユーザーは自覚していないというところに問題があるともいわれている。

そして、この方式は、検索結果のリストの上位にあがった情報が、入力されたキーワードの正しい検索結果であるとは理論的に必ずしも言えないということである。グーグル自身も明確にそのことを認め、表明している。

(注¹：「グーグルの検索結果は、入力したキーワードとの関連度の強さをコンピュータが判定して、結果表示の順位が決まる仕組みだ。これは情報の正しさとしては無関係で、極端に言えば検索結果のトップに表示された情報が「真っ赤なウソ」ということもある」(グーグル日本法人社長村上 憲郎 日本経済新聞 2007年4月12日))

いずれにしても現在の「キーワード型検索システム」には、検索結果とそこから導かれた情報をどう理解したらよいか、あるいは検索結果にでてこなかった情報をどう考えるか、大きな疑問と課題がある。

グーグルのシステムは「群衆の叡智」で、経済学では新古典学派の考えに似ている。新古典学は、経済の構成要素は砂漠の砂のごとくそれぞれ独立して、無関係なものが市場で自由に競争するなかで価格の媒介で経済活動が進むと考える。この新古典学派は市場原理主義をもたらし、経済社会を混乱する弊害があることが再認識されつつあるところであることは言うまでもない。一方制度派経済学は経済の構成要素はいろいろの意味で互いに関連しているのであり、その関連の紐付けを通して経済活動が進展してゆくと見る。これはオントロジーの考えを基にしたものであるといえる。いま新古典経済学派が見直されている所以である。

(注²：グーグルは、一般検索の結果について

は、まさに「美人投票」で、結果の正しさには関心が無く、その美人投票のクリックの多さにお金を払う人がいることを中心にしてビジネスモデルをつくっており、多くの人が集まるようにこのサイトの魅力をつけるために、無料の Google Map, Google Earth, Google Analytics, Google Desktop, Google AdWords, Google AdSense などの餌をつけている)

(注³:「PageRank は、Web の膨大なリンク構造を用いて、その特性を生かします。ページ A からページ B へのリンクをページ A によるページ B への支持投票とみなし、Google はこの投票数によりそのページの重要性を判断します。」(グーグルの Home Page から))

(注⁴:グーグルの検索結果の表示ランクは自動でやっているということであるが、実際は「グーグルは、自らが握っているランキングを自由に決める権利を行使して、何がスパムで何がそうでないかを検索結果から削除によってしめしている」(『グーグル革命の衝撃』NHK 取材班 NHK 出版 P.160),「グーグルのランキングは、数学的

に厳密に運用されてはいるが、あくまでも人がアルゴリズムを決めており、それはグーグルの単なる意見であるという」(同書 P.172)つまりグーグルのサイトでは検索結果の表示において人為の操作が可能であるということである。しかし、グーグルの「検索連動型広告」という仕組みとビジネスモデルは、旧来の広告収入市場をウェブ上に引きずりこんできたという意味では、最近の最も大きなイノベーションといえよう。)

現在のウェブ上での検索システムで、Query Pairs での検索結果を調査した資料がある。

(表-1)は、現在のキーワード型検索システムで具体的にどんな検索結果がでているかを調査した結果である。この調査は、二つのキーワードをペアーにして、この二つの違ったキーワードを入力すると同じ結果がでてこなければならないものである。しかし現在の検索サイトでは、(表-1)のように、極めて精度の悪いものであることがわかる。二つのキーワードは基本的には意味的には同じものであるために、同じ検索結果が得られなければならないはずであるが、それがそうなって

(表-1) Study of Five Asian B 2 C Yellow Pages Sites and Two Tier One Search Engines Local Search
Query Pairs Analysis
Study performed on Thursday, November 2, 2006 and Friday, Nov. 3, 2006 Confidential Information of WAND, Inc.

	Site One	Site Two	Site Three	Site Four	Site Five	Site Six	Site Seven
	Correct results %						
Query Recognition results are returned (accuracy of results not confirmed)	50.00%	70.41%	44.90%	91.84%	13.78%	92.86%	97.45%
SYNONYM PAIRS example: "cars" and "automobiles"	1.28%	0.00%	0.00%	0.00%	1.28%	5.13%	100.00%
SINGULAR / PLURAL example: "book" and "books"	33.33%	100.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SPACING example: "web site design" and "website designer"	0.00%	0.00%	0.00%	0.00%	16.67%	16.67%	37.50%
MORPHOLOGIC VARIATIONS example: "nurses" and "nursing"	12.50%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
COMMON MISSPELLINGS example: "accommodation" and "accomodation"	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

いないという結果である。これが検索者のフラストレーションとなっている。

4-3) 情報検索における基本的な問題点

ウェブ上での言葉による情報検索には基本的な問題点がある。

一つは、ピーター・ロジェのいう「言語の弾力性」の問題であり、これは、情報量が多くなればなるほどの確かな情報の検索が困難になるということである。的確な情報を検索しようとすれば、膨大な無関係な情報を検索しなければならず、これがもっと危険なことであるが、関連情報を見逃すことになる。

もう一つは、通常のコンピュータでの検索では、「リコールと精度の二律背反」という問題もある。「リコール」(総適格項目数に対する適格項目の検索結果数の割合)と「精度」(総検索結果数に対する適格項目の検索結果数の割合)が二律背反の関係にあることである。精度を高めようとすると求める適格項目が検索結果としてもれてきて、検索結果数を高めようとすると精度が落ちるというジレンマに陥ることになる。

つまり、ますます自己増殖する膨大な「非特定の情報」(Nonspecific information)から自分の欲しいもの、知りたいものを検索するのは大変困難になってきており、検索者のフラストレーションが高じてきている。

現在の検索システムは「キーワード検索」で、探したいものに対して、それを意味するキーワードを入れて検索しても、「見つかりませんでした」「そのようなものは存在しませんでした」という検索結果の回答か、あるいはその「キーワード」によって検索されたものが、何百、何千とでてきて検索者はどれが正解か判断がつかないものが多い。インプットしたキーワードが間違いか、適切でないものだったのか、検索者には見当が付かない。

キーワード検索において、こうした問題がでてくるのは、今日の言語の性質、内容の根本に起因する。

その根本とは、人間の脳での言葉による思考のあり方である。人間が頭脳で思考しながら行動しているのは、認識を対象とするいろいろのものが言葉を通じて総て何らかの意味でそして形で関連し、繋がっているからである。つまり人間の生きる世界の森羅万象がことごとく関連しているからである。

Six Degrees of Separation という理論がある。ランダムに選んだ2人の人の関係は、一人の人の関係する他の人を6人辿ると、もう一人の人にたどり着くという。これは、ハーバードのStanley Milgramが1967年に実験したことであるが、実験結果では、5.5人の人で関係をたどると最初の2人が互いに関係があることが分かる。それほど世界の総ての人が繋がっているということである。

人間の脳による思考を考えると、人間の脳の構造は、100億に上る脳細胞を膨大な配線としてのシナプシスにより複雑なネットワークが形成されている。シナプシスは人間が思考し、行動するたびにそれを記憶するために脳細胞を繋ぎ、またそれが確かなものになるとシナプシスの太さが太くなるという構造をもつという。人間の脳は100億以上の脳細胞からなり、その脳細胞に情報が蓄積されるが、それのみでは人間は知的思考をすることはできない。その人間の無限の思考は、いろいろとネットワークとして張り巡らされたシナプシスによっていろいろの脳細胞に記憶された情報を引き出しながら行われるのである。

つまり、人間の思考はいろいろな事象の記憶された脳細胞を手繰りながら、あるものを思考し、認識をするが、それは脳細胞に記憶されているいろいろの知識、データをシナプシスで繋げ、手繰りながら行うのである。知識・情報が脳のなかで構造化されているのである。そのシナプシスが切れていれば、孤立した知識を記憶した脳細胞は何も人間の思考に寄与しないといわれている。

より知的な活動ができるかは脳細胞にある記憶された膨大な情報が互いに如何に多くのシナプシスにより連結されているかによる。

このシナプシスによるネットワーク回路の数は、

順列組み合わせの数で、宇宙に存在する分子の総数よりも多いといわれるほど膨大な数に上る。つまり、人間の意識、思考のプロセスでは、認識としての言葉はそれぞれ孤立しているのではなく、丁度脳細胞とシナプシスのネットワーク回路のように、それぞれ互いに意味的に繋がっているのであり、それにより思考ができる。

こうした脳の仕組みと同じように、人間の使うどんな概念、言葉、イメージも、実質的な多くの関連を通じて、他の言葉、概念、イメージと複雑な繋がり、関係（インターコネクト）をもっているものである。コンピュータを使った知識の検索においては、検索対象の言葉を脳が働くのと同じように、総ての言葉を意味的に関連付けるような構造化が必要になる。さもないとコンピュータは人間の思考を助けてはくれない。したがって人間の脳の中でのそれぞれの言葉は孤立しているのではない。お互いに意味的に繋がっているのである。しかし残念ながら、コンピュータでは、言葉は、何もしなければ、繋がっていない。

この言葉のインターコネクトが、人間の頭脳が働く基本であり、情報化社会の検索システムにおいても、言葉の「インターコネクト」がキーとなり、これが前述の「リコールと精度のディレンマ」の問題を解決してくれるものである。

ウェブ上の検索システムの基本構造は、これまで述べてきたことから分かるように、「検索エンジン」と「被検索対象の言葉のデータベース」という二つの構造からなる。これまで「検索エンジン」の高性能化が進められてきた。しかし検索の対象になる言語のデータベースのきちんとした分類構造が、エンジンの性能以上に重要なものであることがやっと認識されてきたのである。検索が人間の思考のように進むには、この被検索対象の言語のデータベースを構造的オントロジ・タクソノミとして構築する必要があるということである。

コンピュータによる検索システムでも、こうした人間の脳が行う思考プロセスと同じような働きをさせるには、言葉のデータベースにある工夫が必要になる。

これからのウェブでの検索システムは、人間が使うのであるから、脳の思考プロセスに沿った、このような構造にならなければ、本当の知識社会は到来しない。

現在の欧米のキーワード型検索システムでは、膨大なタグをつけたキーワードを持っているが、それらはほとんどランダムにデータベースの箱の中にはおり込まれている状態で、人間の脳細胞とシナプシスのような構造的に整理されていないことに問題がある。

自動車の例を使って比喻で表現すると、セルフ・スターターやオートマチック・トランスミッションが発明される以前の自動車が、今日のウェブ上での検索システムであるということが出来るであろう。エンストやエンジン始動が困難で、ギアシフトでトラブルが起こるようなものである。

しかし今日のような運転しやすい自動車になるのは時間の問題である。

4-4) これからの検索システムの方向

以上のように情報とは、人間、生物が世界の森羅万象と関係することで出現するものであり、その意味では言葉としての情報は、あらゆるものが意味的に繋がっていると考えなければならない。つまり「言葉のシンデティック」(Syndetic)と呼ばれているものであるが、それは、情報科学においては一緒に括る、あるいは連結するという意味であるが、語彙のクロス・リファレンス(参照)を通じて概念のリンクを作り上げることである。その標準的な手法は、情報科学において、構造化された語彙 (Controlled Vocabularies) を使ってあらゆる概念の間の関係を作り上げることである。これを「構造化された Taxonomy」、あるいは「構造化された Ontology」とも呼ぶ。

この構造には、階層性、等価性、関連性というものがある。あるドメインにおいて言語はより上位の、より抽象的なものから、具体的な言語という関係が存在する、また同じ意味するものでも違う言葉、文字が存在する。あるいはある意味に関連した違った言葉がある。これらを意味的に関連

付け、紐をつけて構造化されたものとして理解しなければならない。人間の脳はそれをシナプスで紐付けているので理解できるが、コンピュータは言語のデータをそれぞれの意味においてリンクしたデータベースを整えておかなければ、それを認識できない。

具体的には基本語 (Preferred Term)、関連語 (Related Terms)、兄弟語 (Sibling Terms)、更には類似語 (Synonyms)、延長語 (Extended Terms)、同形異義語 (Homographs)、また異表記語 (Orthographic Term)、がある。

更に分類データの中には、ものごとの属性 (Attribute) と属性値 (Attribute value) というデータベースがある。これがあると求めるものの具体的なスペックを考慮した最終のものが検索、獲得できる。

グーグルをはじめとする現在の欧米のキーワード型検索システムのキーワードはこうした構造化はされていない。グーグルは PageRank という検索連動型広告というビジネスモデルを採用しているために、むしろ構造化を避けているといえよう。グーグルの検索サイトで、データベースを構造化をすれば、これまでの PageRank というビジネスモデルが崩壊する。

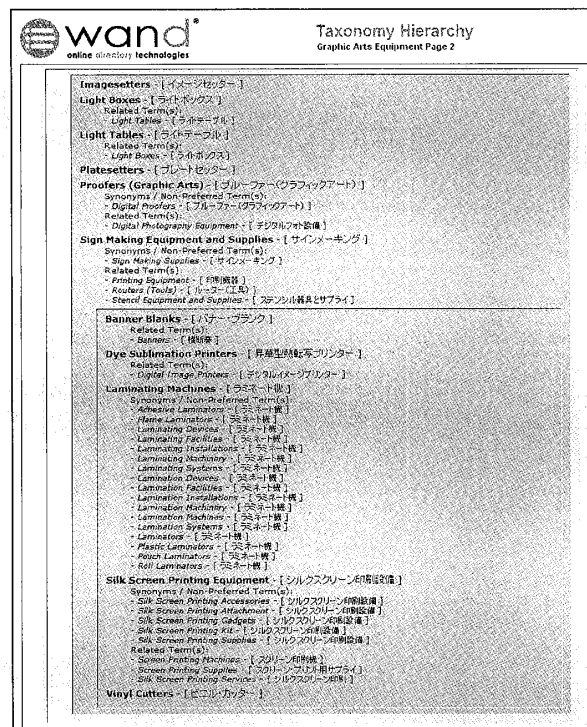
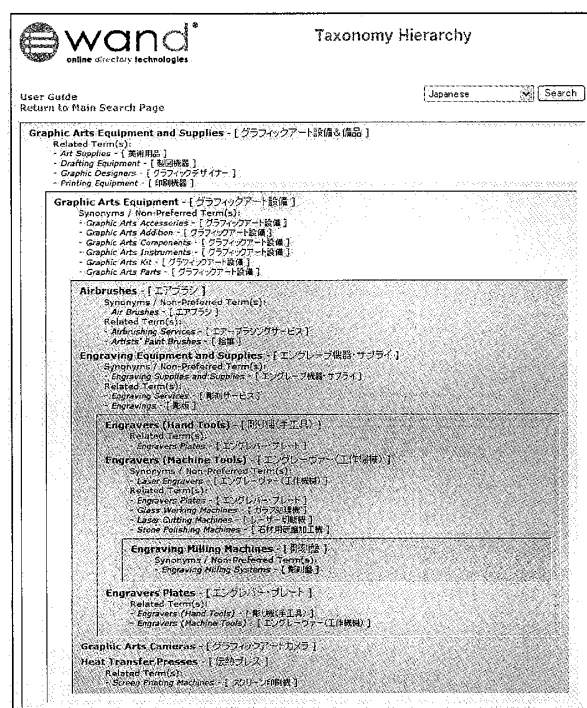
構造化された Taxonomy のデータベースの例が、次の (表-2) である。

(アメリカの Wand 社のもの)

(注⁵: Wand 社の Taxonomy は2007年6月現在において、68,000以上の製品およびサービスの項目と1,400,000以上の属性と属性値をもっている)

例えば、「眼鏡 (メガネ)」という言葉を取り上げると、「眼がね」、「眼鏡 (ガンキョウ)」、「メガネ」、「アイグラス」、「アイウエア」、「スペクタクル」、「サングラス」、「拡大鏡」というような類似語、関連語があり、また「頑強 (ガンキョウ)」、「妻がね (メガネ)」という同形異義語がある。もちろん綴りを間違えた異表記語は沢山ある。これにはさらにスキーの「ゴウグル」、「眼科」、「ファッションメガネ」などという「関連語」、「兄

(表-2)



wand
online creative technologies

Taxonomy Hierarchy
Graphic Arts Equipment Page 4

Graphic Arts Supplies - [グラフィックアート用品]

Related Term(s):

- Art Supplies - [美術用品]
- Software - [ソフトウェア]

Graphic Arts Chemicals - [グラフィックアート化学製品]

Blanket/Roller Washes - [ブランケット/ローラーウォッシュ]

Fountain Admixes - [噴水用添加剤]

Fountain Solutions - [噴水用溶液]

Synonyms / Non-Preferred Term(s):

Fountain Applications - [噴水塗料]

Fountain Preps - [噴水剤]

Fountain Solvents - [噴水溶剤]

Fountain Systems - [噴水装置]

Hobby Knives - [ホビーナイフ]

Synonyms / Non-Preferred Term(s):

Exact Knives - [エクザクナイフ]

Hobby Cutters - [ホビーカッター]

Accuracy Knives - [エクザクナイフ]

Exact Knives - [エクザクナイフ]

Litho Varnishes - [リトグラフ用ニス]

Related Term(s):

Varnishes - [ニス]

Premask Tape - [プレマスク用テープ]

Synonyms / Non-Preferred Term(s):

Graphic Arts Tape - [グラフィックアートテープ]

Screen Printing Supplies - [スクリーンプリント用材料用品]

Synonyms / Non-Preferred Term(s):

Screen Printing Supplies - [スクリーンプリント用材料用品]

Screen Printing Ink - [スクリーンプリント用インク]

Related Term(s):

Silk screen printing equipment - [シルクスクリーン印刷装置]

Scoop Coaters - [スクープコーター]

Synonyms / Non-Preferred Term(s):

Screen Printing Emulsions - [スクリーンプリント用感光剤]

Screen Printing Frames - [スクリーンプリント用フレーム]

Synonyms / Non-Preferred Term(s):

Screen Printing Squeegees - [スクリーンプリント用スクイジー]

Wire Cloth - [ワイヤークロス]

Synonyms / Non-Preferred Term(s):

Wire Cloth - [ワイヤークロス]

Screen Printing Meshes - [スクリーンプリント用メッシュ]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Wire Cloth - [ワイヤークロス]

Stripping Supplies (Graphic Arts) - [スクリーニング用品(グラフィックアート)]

Layout Film - [レイアウトフィルム]

Masking Film - [マスクングフィルム]

Related Term(s):

Striping Film - [ストリピングフィルム]

Masking Paper - [マスクングペーパー]

Stripping Base - [ストリピングベース]

Related Term(s):

Masking Film - [マスクングフィルム]

にも、現在の欧米のキーワード検索システムのように、データベースが意味的に構造化されていないものとそれが Ontology 的に構造化されているものの二つものがある。

その二つの検索システムのプロセスの違いを見てみよう。

「構造的オントロジ・タクソノミをもたないもの」

- 1) 情報をインデキシング（タグをつける）する。
- 2) 検索窓でキーワードを入れる。
- 3) そのキーワードでインデキシングされたキーワードを検索して、その意味を考えないで、インデキシングされたキーワードと文字的に同じものをピックアップする。（キーワードと意味的に関連あるものがピックアップされるとは限らない）
- 4) そのピックアップされたキーワード情報を特殊なアルゴリズムでランク付け、しかもそのキーワードに繋がった情報をリンクさせる。そこで検索連動型広告システムで広告記事を添付させる。
- 5) これを検索結果として表示する。

この検索結果は、前述のとおり、キーワードの意味的理解はしていないので、文字的にキーワードと一致している情報で内容的に関係のないものも検索結果として表示してしまうことになる。現在の欧米の検索システムはこのタイプである。

「構造的オントロジ・タクソノミをもったもの」

- 1) 情報を意味的関連においてインデキシングし、「構造的分類（Ontology Taxonomy）データベース」にマッピングする。同時に新しい言語、新しいコンセプトのものが出現した場合、人間が判断しそのものを「構造的オントロジ・タクソノミ・データベース」の中に、意味的に的確に位置づける。
- 2) 検索窓でキーワードを入れる。
- 3) 「構造的オントロジ・タクソノミ・データベース」から「意味的に」そのキーワードに該当するものをピックアップする。

- 4) それをグルーピングして、検索結果として表示し、それに関連する情報をリンクさせる。

これに加えて属性と属性値をキーワードとあわせて検索窓に入れる設定をした場合、「構造的分類、属性・属性値データベース」より求める商品、サービスをピンポイントで検索できる。この構造的オントロジ・タクソノミを持った検索システムがこれからの検索システムであるといえる。つまり Semantic Search（意味づけされた検索）である。

特に日本語はもともと中国を含めて外来語が多く受け入れられ、発展してきたのであるが、近年、歴史の浅い欧米の外来語がカタカナとして日本語の中にどんどん入ってきている。これ外来語の意味が、欧米で使われている意味とは違ったものとして日本で理解され、外国人とのコミュニケーションで誤解を生んであるケースが多々あるといわれている。

こうした意味でもそれぞれの言葉の意味を中心に、その関連性を紐づけることは大変重要になる。

つまりこれは検索キーワードの「意味を理解して」、それにフィットする情報を取り出すものである。グーグル検索システムとは違った。こうした検索キーワードの意味を理解することを前提に「自然言語」での検索システムの開発があちこちで進んでいる。（例：Powerset 社など）

ところで、人間の世界は一定不変ではなく、常に変化している。常に新しい技術、製品、新しい社会、社会制度、新しい言葉の意味が間断なくつくり出され、変化している。この変化は同時に、こうした森羅万象を映し出す構造化されたオントロジ・タクソノミの内容が変化しなければならない。

それではその変化する世界をどのようにデータベースに常にアップデートしていくかが問題になる。グーグルは検索の結果を出すのは人間の手ではなく自動でやるといっているが、結果を出すアルゴリズムは常に変えているといっている。しかしこれは実質的には、グーグルの結果を導き出す仕組みは人間が操作しているということになる。

つまりこの構造的オントロジ・タクソノミ・データベースは、社会の変化によってそれをアップデートしていかなければならないが、新しいコンセプト、新しい言葉、新しい言葉の使い方が出現した場合、クローラーで自動的に収集し、整理するが、その新しい言葉を構造的なオントロジ・タクソノミのどこにマッピングするか、構造を変更するかは、これは現在のところコンピュータで自動的に処理することはできない。これは人間の判断によるしかない。

(注⁶：Wand 社の場合、250人ぐらいの世界のいろいろのドメインエキスパートをネットワークでかかえており、新しいコンセプト、言葉がでてきたらこれ等のエキスパートが最終的に判断して、タクソノミ・データベースに組み入れる。)

(5) 現在のキーワード型検索システムを超える仕組み

創発的な思考のプロセスがイノベーションには必要であるが、これをウェブでの検索の場でももてたという声はずっとでていた。

現在のキーワード型検索システムでは、ウェブから適当な情報、記事をピックアップし、「コピペ」(コピー&ペースト)して簡単に論文や提案書を作成する傾向があるが、各「情報小包」をコピーし、ペーストしただけでは中身がなく、矛盾するものもあり、言葉の遊びになる。このコピペは、ものの本質、おのおのの事柄の意味的つながりを理解、究明しないで、つぎはぎして語呂合わせの空虚な論理を構成することの危険性が心配されている。ある事柄の判断材料も、自分自身の記憶もすべて検索の窓のむこうがわに盲目的にゆだねるという弊害を多くの人が憂いている。思考停止が起こる。

また検索連動型の広告は、個人の無意識の好みに働きかけて、情報を大衆に与える方法は個人の判断を誘導することになる危険性があるといわれている。

こうした危険性は現在のキーワード型検索シス

テムの構造から来ている。

キーワード型検索システムでタクソノミが構造化させていないものは、つまり現在の欧米の企業がドライブする検索システムは『智』ではないといわれる所以である。

比喩でいえば、「キーワード型検索システム」(構造化されていないもの)は、魚の一本釣りのようなものである。魚釣りの人には、水中にどんな魚がいるか可視化できない。そこにこんな魚がいるであろうという推定のもとに、釣りをするのである。ある「キーワード」という意味で、特定の釣り針と餌をつけた一本釣りでは、その針と餌にあう魚しか釣れない。それもその魚がそこにいればの話である。あるいは同じ釣り針と餌でも、その取り付け方の違いによってはその魚がいても釣れないということが起こる。場合によっては魚でなく藻やゴミがつりあがることになる。最初に目指したある特定の魚以外の欲しくなるような魚がいても、特定の針、餌というキーワードではそれが釣れない。

しかしこれでは、知識社会での「知識創造」「イノベーション」はできない。新しい情報の検索システムが必要である。まず、検索被対象のものが検索者にとって、構造化され整理された形で、可視化され、トランスパレントになることである。瓦礫の山がトランスパレントになっても意味がない。そこで更に必要なことは、確実に自分の期待したものが得られると同時に、思考のプロセスで新しい発見、飛躍が出来る知識情報の検索の仕組みが必要である。

あるドメインにおいてダイバーが水中にもぐり、きちんと構造化された Taxonomy で整理されて情報、書類が並んでいて、それが検索者に明確に可視できるものを採るという仕組みが、これから構築されるべきシステムである。そこには検索の最初の関心事以外の発見、創発的な飛躍が得られるのである。アワビをとろうと思って潜ったが、そこに立派な蛸がいて、これをとることもできる。これがこれからの知識社会での「ソフト・バスでの知識活動」である。

つまり現在存在する検索システムは、情報の検索の場を提供する側から押し付けるといふ発想であり、最近、ワン・ツー・ワン・マーケティングのシステムで、ある個人の性格、行動、考えという個人情報を集めて分析し、それに基づきその人が好みそうな商品の広告を送りつけるという仕組みが開発されつつあるが、これはやはり情報検索の場を提供する側からの情報のお仕着せであり、これは個人情報の問題からいずれ禁止されるものと見られている。

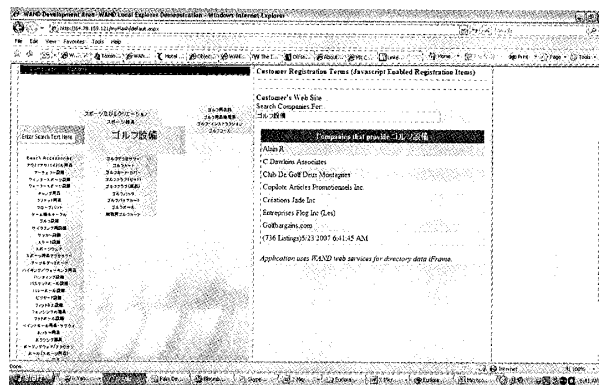
そうではなくつまり他からのお仕着せではなく、国民大衆の個人が自分で欲しいものを探せる場を提供することが必要である。

換言すれば、これまでの情報は大企業の独占であったが、ウェブ、インターネットにより情報はオープン化してきたが、しかし現在はその情報のオープン化でも情報の検索サイトの提供者からのお仕着せであったものを、これからは検索者の主導による情報の獲得の時代を創る必要がある。

水中ダイバーのように水中に存在するあるドメインの情報、知識が意味的に紐付けられ、構造化されてたものが明確に可視化されていけば、その人の判断により、いろいろの有用なものが発見され、新しいイノベーションに繋がっていくことになる。

人間の脳の働きとウェブに関して、茂木健一郎氏は興味ある見方をしている。「脳とインターネットの関係は大変おもしろいテーマです。インターネットの世界は、私の言葉でいう『偶有性』、つまり、ある事象が半ば偶然的に半ば必然的に起こるという不確実な性質に充ち満ちています。…スモールワールド・ネットワーク性（一見遠く離れているものどうしも、少数のノードを通して結ばれている性質）という観点からみると、そっくりです。…脳とコンピュータのちがいというのは、チューリング・マシンをプログラミングするというのはいわば『ウェブ1.0』で、脳は要するに『2.0』以降です。』（『フューチャリスト宣言』梅田望夫・茂木健一郎 ちくま新書 P.28, 29）

アメリカの Wand 社は Explorer というコンセ



プトで新しいサイトを計画している。

例えば「ゴルフ用品」「ゴルフ設備」というキーワードでいったんあるドメイン（領域）に入ると、もはやキーワードの入力はいらぬ。そのドメインの中には、前述したように階層の抽象度の高いものから低い言葉、類似語、関連語、兄弟語が意味的にすべて関連付けられて設置されており、ゴルフシューズ、ゴルフアパレル、ゴルフ用品の、アイアン、ドライバー、パター、サンドウエッジ、あるいはエンターテインメントとしてのゴルフクラブ、カントリークラブ、ゴルフトーナメント、ゴルフ教師、それから更に兄弟語として、テニス教師、テニスシューズ、テニス用品へと飛んでゆける。

化学製品のドメインでナイロンにたどり着くと、ナイロンを使用しているいろいろのアプリケーションに展開でき、兄弟語としていろいろの分野のナイロンの応用製品の事例をサーフすることができる。こうしたサイトで新しい発想、新しい創発が起こり、イノベーションに繋がることになる。

この中にリンクでウイキペディアのような百科事典のようなサイトに飛ばして、そのものの更に詳しい内容を理解するということも出来る。ビジネスとしてもゴルフ用品をクリックすると左の画面にそれを供給するサプライヤーが出てきてそこで広告収入のビジネスが生まれる。

「連想検索」というコンセプトがある。「私たちの頭の中には大量の情報が蓄積されていて、情報を与えられると、関連がありそうな情報が呼び起こされる。同じように電脳空間の情報プールの中からざっくりと近そうな情報を集めてきて定時

すると、それがまた人間の脳を刺激して脳内で新たな情報を呼び起こす。そういった人間と機械の両方の連想の相互作用が知的な作業のために重要である」

アメリカのベンチャーのスタンブル・アボン社がこうしたシステムに挑戦している。

つまり、これからの知識社会として人間が自分の求める、欲する情報、知識、データを膨大な無秩序にウェブ上に存在するものから、公正に、的確に引き出す仕組みがなければならない。

同時にこれから知識を、情報をもとに新しい社会に必要な商品を創るパラダイムシフトをもたらすようなイノベーションを進めるには、単に自分が欲しいものを「キーワード」をインプットして探すということから、自分が気がついていない、意外なものを発見するプロセスが望まれているのである。つまりキーワードとしては明確にインプットできない何かを探すことがこれから多く起こることになる。人間の思考を促進し、新しいものの発見をサポートする仕組みが望まれているのである。

一つ問題があるのは、21世紀に入り世界はグローバル化の波の中で、すべてが物凄い勢いで変化をしていることである。とくに情報がウェブの世界にオープン化され、世界のあらゆる人が情報、知識の加工に参画したフラットな社会になり、科学の進歩ともあいまって新しい知識、言葉がどんどん加わってきている。そうすると被検索対象のデータベースも新しいものをその構造化された意味づけの網の目のなかに正しく配置しなければならない。属性も従って変化してくる。そうなると固定的な構造化されたデータベースはかえって機能しなくなる。さりとてランダムに瓦礫の山のようにデータを積み上げることも出来ない。しかも新しいキーワードを意味的にどこに関連させ、どこに配置させるべきかは、現在のところコンピュータでは判断できない。ここはまだ人間の叡智の働く場所である。Wand社は各ドメインの世界の専門家を契約ベースでネットワークで組織しており、新しいキーワード、分類タームがサイト

から指摘されるか、あるいはウェブロボット（クローラー）により採集されたものを、発見後5ワーキング・デー以内にそれを決定している。こうした専門家のネットワークが必要である。

知識社会におけるソフトバスの形態は現在の欧米のリードする検索システムのみでは不十分である。グーグルはウェブの広告という意味では新しいビジネスモデルを築きあげたが、前述のとおり、正しい検索結果という意味では問題と制約があるので、別の形態があってしかるべきである。

これまでの欧米の企業がドライブする検索サイトのビジネスモデルは、最初は「Pay for Banner」広告をバーナーで流すという方法をとった、しかしグーグルが検索連動型広告のビジネスモデルを広めた。つまり「Pay for Link or Click (Page rank)」である。しかし、これからはユーザーの求めるものが得られるところで支払い、広告収入が起るようなビジネスモデルも出てくるはずである。つまり「Pay for Relevant Result」である。

いずれにしても新しい知識創造検索システムの構築には「構造的オントロジ・タキシノミ」のデータベースを開発整備することが必要である。言語としては日本語は構造的に英語よりも困難であり、しかも日本にはそうしたデータベースは断片的にいろいろの企業、組織に散在するが、共通のデータベースになっていないことが問題である。

こうしたデータベースは現在のところ一企業では出来そうも無い。これを国家プロジェクトとして早急に開発整備する必要がある。グーグルのマネをしても仕方が無い。それを超えるものが必要で、しかも、各国の国家保安のために、世界に複数個の独立したデータベースをもつ検索サイトが必要であることを認識しなければならない。

この「構造的オントロジ・タキシノミ・データベース」は、これから展開してくる多言語の「自動音声認識、翻訳通訳システム」の基礎技術、基礎データベースとなるものである。

まとめ

前述のように21世紀の知識社会では、商品の主とした付加価値は知識、情報により創造させることになる。この新しい商品の開発により多くの人々が参画するには「ソフト・バス」という概念の知識・情報のインフラが構築されなければならない。その知識・情報のインフラとして重要なものの一つがウェブにおけるパブリックに開けた知識検索

システムである。これは今日の欧米の企業の検索システムではその目的を果たさない。膨大な細分化された知識、情報を意味的に構造化したものをベースにしなければならない。

しかも単純なキーワード型検索ではなく知識の創発を起こすような検索システムが必要になる。この優れたものを経済社会の基盤としてもつことがその国、地域の国力となる。その具体的な道を示した。